Simulation-guided rational *de novo* design of a small pore-forming antimicrobial peptide

Charles H. Chen[1,2,3,*], Charles G. Starr[4], Evan Troendle[1,2,3], Gregory Wiedman[5], William C. Wimley[4], Jakob P. Ulmschneider[6,*] Martin B. Ulmschneider[1,2,3,*]

[1]Department of Chemistry, King's College London, London, UK
[2]Department of Materials Science and Engineering, Johns Hopkins University, Baltimore, MD, USA
[3]Institute for NanoBioTechnology, Johns Hopkins University, Baltimore, MD, USA
[4]Department of Biochemistry and Molecular Biology, Tulane University, New Orleans, LA, USA
[5]Department of Chemistry and Biochemistry, Seton Hall University, South Orange, NJ, USA
[6]Institute of Natural Sciences, Shanghai Jiao-Tong University, Shanghai, China

ABSTRACT

In the age of failing small-molecule antibiotics tapping the near-infinite structural and chemical repertoire of antimicrobial peptides (AMPs) offers one of the most promising routes towards developing next-generation antibacterial compounds. One of the key impediments *en route* is the lack of methodologies for systematic rational design and optimization of new AMPs. Here we present a new simulation-guided rational design approach and apply it to develop a potent new AMP. We show that unbiased atomic detail molecular dynamics (MD) simulations are able to predict structures formed by evolving peptide designs enabling structure-based rational fine-tuning of functional properties. Starting from a 14-residue poly-leucine template we demonstrate the design of a minimalistic potent new AMP. Consisting of only four types of amino acids (LDKA), this peptide forms large pores in microbial membranes at very low peptide-to-lipid ratios (1:1000) and exhibits low micromolar activity against common Gram-positive and Gram-negative pathogenic bacteria. Remarkably, the four amino acids were sufficient to encode preferential poration of bacterial membranes with negligible damage to red blood cells at bactericidal concentrations. As the sequence is too short to span cellular membranes, pores are formed by stacking of channels in each bilayer leaflet.

Keywords: Antimicrobial Peptides, Polyleucine Peptide, LDKA, Molecular Dynamics, Protein Folding, and Pore-formation.

INTRODUCTION

Pore-forming antimicrobial peptides (AMPs) are powerful and ubiquitous components of the innate immune defence in all domains of life. AMPs are amphiphilic peptides that selectively target and kill a wide variety of microbial pathogens at low micro-molar concentrations.[1-3] Despite the discovery of thousands of AMPs over the last 100 years,[4] the molecular mechanisms driving antimicrobial activity remain poorly understood.[5] Known AMPs vary widely in number of residues, sequence, and secondary structure, and no common activity motif has been discovered to date. This lack of sequence-function relationship is unusual for proteins and has hindered efforts aimed at understanding the root causes of AMP activity. For example, a few amino acid mutations in melittin can result in dramatic changes of pore stability,[6] potency,[7-9] pH-sensitivity,[10, 11] and selectivity for particular membrane types.[8] This presents a significant challenge for rational *de novo* design of AMPs.

At present, designing pore-forming AMPs can in principle be performed using helical wheel projections, assuming simple channel structures where parallel oligomeric helix bundles surround a central aqueous pore. In general, channel-forming AMPs are designed by empirical modification of peptide amphiphilicity and variation of the hydrophobic and hydrophilic faces of the helix using peptide libraries. Subsequent high-throughput screening is used to select functional sequences; however, the structure and function of AMPs in fluid membranes are dynamic and transient, and a single peptide sequence can result in several different oligomeric structures, that typically deviate significantly from simple symmetric helical assemblies.[12] This suggests that successful rational design of AMPs requires knowledge of the dynamic structural ensemble in the membrane, rather than a static structural model.

Here, we report a new methodology that uses unbiased atomic detail folding-partitioning MD simulations to guide rational *de novo* AMP design. We have previously shown that MD simulations in the multi-microsecond regime are sufficient to capture peptide folding, insertion, and spontaneous oligomerisation into structural assemblies in the bilayer at atomic detail.[12-18] In this approach, fully flexible peptides are allowed to freely assemble and dissociate in the bilayer, thus revealing the structural ensemble present at equilibrium.[19] Like an equivalent laboratory experiment, this approach is completely unbiased. The method yields the relative populations and lifetimes of individual peptide assemblies, which directly relate to structural stability.

A single simulation of ~5 μs is usually sufficient to capture the formation of an ensemble of structural assemblies in the membrane. Simulations of this type are now feasible on modern hardware. For example, the 5 μs spontaneous membrane assembly simulations performed in this paper typically take 2 months on a modern GPU. Analysis of the architectures of the peptide assemblies formed during the simulation, and their relative populations, provides atomic detail insights into interactions that are important for structure formation and stability. We show here that this information can be used to tune peptide pore-formation properties rationally.

We start with a 14-residue polyleucine sequence that is progressively functionalised to develop a soluble amphiphilic peptide that forms membrane pores. This simulation-guided design and refinement process, which utilises atomic-detail dynamic structural and functional information obtained from MD simulations, ultimately results in the discovery of a powerful pore-forming AMP sequence.

MATERIALS AND METHODS

Peptide synthesis. LDKA peptides (with free termini) were synthesized using standard Fmoc chemistry and purified to 98 % purity using reverse phase HPLC. Peptide purity and identity were confirmed by HPLC and ESI mass spectrometry.

Large unilamellar vesicle (LUV) preparation. 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG), and hexadecanoyl sphingomyelin (Egg SM; PSM) were purchased from Avanti Polar Lipids. Lipids were dissolved in chloroform, mixed, and dried under nitrogen gas in a glass vial. Any remaining chloroform was removed under vacuum overnight. To make LUVs lipids were resuspended in 10 mM sodium phosphate buffer (pH = 7) with 100 mM potassium chloride. LUVs were generated by extruding the lipid suspension 10 times through 0.1 µm nucleopore polycarbonate filters to give LUVs of 100 nm diameter.

ANTS/DPX leakage assay. 5 mM ANTS and 12.5 mM DPX were entrapped in 0.1 µm diameter extruded vesicles with lipids. Gel filtration chromatography of Sephadex G-100 (GE Healthcare Life Sciences Inc) was used to remove external free ANTS/DPX from LUVs with entrapped contents. LUVs were diluted to 0.5 mM and used to measure the leakage activity by addition of aliquots of LDKA. Leakage was measured after 3 h incubation. 10% Triton was used as the positive control to measure the maximum leakage of the vesicle. Fluorescence emission spectra were recorded using excitation and emission wavelength of 350 nm and 510 nm for ANTS/DPX using a BioTek Synergy H1 Hybrid Multi-Mode Reader.

Macromolecule release assay. Several different size dextrans were prepared and labelled with both TAMRA and biotin. Conjugated dextran was entrapped in POPC LUVs as described above. External dextran was removed by incubation with immobilized streptavidin. Streptavidin labelled with an Alexa-488 fluorophore was added during the leakage experiment with the peptide as previous described.[6] The sample was incubated for 3 hours before measuring Alexa-488 fluorescence. A control without added peptide served as the 0% leakage signal and addition of 0.05% vol. Triton X-100 was used to determine 100% leakage.

Circular dichroism spectroscopy. LDKA solutions (50 µM) in 10 mM phosphate buffer (pH 7.0) were co-incubated with 800 µM POPC:POPG (1:1) and POPC:CHOL (7:3) LUVs in identical buffer (see *LUV preparation* above). CD spectra were recorded using the synchrotron radiation circular dichroism beamline on ASTRID at Aarhus University. Spectra were recorded from 270 to 170 nm with a step size of $\Delta\lambda$ = 0.5 nm, a bandwidth of 0.5 nm, and a dwell time of 2 s. Each spectrum was averaged over 3 repeat scans. The averaged spectra were normalized to molar ellipticity per residue. The raw data were analyzed using DichroWeb <http://dichroweb.cryst.bbk.ac.uk/>.

Bacterial minimum inhibitory concentration. *Escherichia coli* strain ATCC 25922, *Staphylococcus aureus* strain ATCC 25923, and *Pseudomonas aeruginosa* strain ATCC PAO1 were used in this study. Overnight cultures were sub-cultured and diluted to an initial bacterial cell density of ~3 x 10^5 colony forming units (CFU) per mL in Lysogeny broth. Cell counts were determined by measuring optical density at 600 nm (OD_{600}), with an optimal sensitivity at OD_{600} = 0.3-0.6 in a 1

cm path-length cuvette. OD_{600} = 1 corresponds to 1.5 × 10^8 CFU/mL for *S. aureus,* 5 × 10^8 CFU/mL for *E. coli*, and 2.04 × 10^8 CFU/mL for *P. aeruginosa*. Bacteria were added to peptide (LDKA and indolicidin) dilutions (1.3, 2.0, 2.9, 4.4, 6.6, 9.9, 14.8, 22.2, 33.3, 50.0, and 75.0 μM) and co-incubated at 37 ˚C. After 12 hr incubation, the optical density of the wells were recorded on a plate reader to determine whether they were sterilized (OD_{600} < 0.08) or were at stationary phase growth (OD_{600} > 0.5). Intermediate values, which were rare, were considered positive for growth. Average minimum sterilizing concentrations were calculated from the lowest peptide concentration that sterilized the bacteria in each serial dilution. The samples were done in sextuplet.

Hemolysis assay. Fresh human red blood cells were obtained from Interstate Blood Bank, Inc., and thoroughly washed in PBS until the supernatant was clear. hRBC concentration was determined using a standard hemocytometer. In hemolysis assays serial dilutions of peptide were prepared, followed by the addition of 2 × 10^8 hRBC/mL. After incubation for 1 hr at 37 ˚C the cells were centrifuged and the released hemoglobin was measured by optical absorbance of the heme group (410 nm). Negative control was buffer only (0% lysis), and the positive control was 20 μM melittin and distilled water (100% lysis). The samples were done in triplicate.

Molecular dynamics simulations and analysis. Unbiased all-atom MD simulations were performed and analyzed using GROMACS 5.0.4 and Hippo BETA simulation packages <http://www.biowerkzeug.com>, and VMD molecular visualization program <http://www.ks.uiuc.edu/Research/vmd/>. The pdb structure of extended peptides (GL_5KL_6G, LDKL, and LDKA) were generated using Hippo BETA (see Table S1, Table S2, and Table S3). These initial structures were relaxed via 200 Monte Carlo steps, with water treated implicitly using a Generalized Born solvent.

After relaxation, the peptides were placed in all atom peptide/lipid/water systems containing model membranes with 100 mM K and Cl ions using CHARMM-GUI <http://www.charmm-gui.org/>. Protein folding simulations were equilibrated for 10 ns with applying position restraints to the peptide. For pore-forming simulations single peptides were allowed to fold onto the bilayer for ~600 ns; subsequently, the systems were multiplied 4 times in both the x and y directions. MD simulations were performed with GROMACS 5.0.4 using the CHARMM36 force field, in conjunction with the TIP3P water model. Electrostatic interactions were computed using PME, and a cutoff of 10 Å was used for van der Waals interactions. Bonds involving hydrogen atoms were constrained using LINCS. The integration time-step was 2 fs and neighbor lists were updated every 5 steps. All simulations were performed in the NPT ensemble, without any restraints or biasing potentials. Water and the protein were each coupled separately to a heat bath with a time constant τ_T = 0.5 ps using velocity rescale temperature coupling. The atmospheric pressure of 1 bar was maintained using weak semi-isotropic pressure coupling with compressibility $\kappa_z = \kappa_{xy} = 4.6 \cdot 10^{-5}$ bar^{-1} and time constant τ_P = 1 ps.

Peptide thermostability enables advanced sampling at high temperatures. The LDKA peptide is resistant to thermal denaturation when bound to the membrane. This allows all simulation to be run at 70 °C (Figure S1), increasing pore-formation kinetics. We have previously demonstrated that elevating the temperature does not change conformational equilibria or partitioning free

energies of helical membrane-active peptides, provided they are stable against thermal denaturation (see Supplement); however, the vast increase in sampling kinetics at high temperatures allows simulation of peptide folding, bilayer partitioning, and pore assembly without the need for advanced sampling techniques that require additional information or may bias the system.[12, 13, 15, 16, 20, 21]

Oligomer population analysis. In order to reveal the most populated pore assemblies during the simulations, a complete list of all oligomers was constructed for each trajectory frame. An oligomer of order *n* was considered any set of *n* peptides that are in mutual contact, defined as a heavy-atom (N, C, O) minimum distance of < 3.5 Å. Frequently, this definition overcounts the oligomeric state due to numerous transient surface bound (S-state) peptides that are only loosely attached to the transmembrane inserted peptides that make up the core of the oligomer. These S-state peptides frequently change position or drift on and off the stable part of the pore. To focus the analysis on true longer-lived TM pores, a cutoff criterion of 65° was introduced for the tilt angle τ of the peptides. Any peptide with $\tau \geq 65°$ was considered in the S-state and removed from the oligomeric analysis. This strategy greatly reduced the noise in the oligomeric clustering algorithm by focusing on the true longer-lived pore structures. Population plots of the occupation percentage of oligomer *n* multiplied by its number of peptides *n*, were then constructed. These reveal how much peptide mass was concentrated in which oligomeric state during the simulation time.

Permutational cluster analysis. Conventional cluster analysis cannot deal with a transient formation process were the same specific oligomeric pore could consist of entirely different peptides each time it forms. In order to address this, a permutational cluster analysis was performed. After the oligomer list was computed, in a 2nd step, all oligomers of the same order *n* were conformationally clustered. A clustering algorithm was used, with a backbone RMSD similarity cutoff criterion of 4 Å. Since each oligomer could be made up of different peptides – or of the same peptides, but in a different order – the clustering had to compare one oligomer with all *n*! permutations of peptide arrangements of the other oligomer. Permutations were generated using Heap's algorithm. The final RMSD value of the conformational similarity was considered the lowest RMSD value as obtained from the *n*! permutational comparisons.

Hydrophobicity analysis of the peptide termini. Terminal hydrophobicities were calculated for AMPs that have 12 or more amino acids in the *APD*. The partitioning free energy was determined for the first six amino acids in the sequence (i.e. at the N-terminus) and the last six amino acids (i.e. at the C-terminus) using the Wimley-White octanol-to-water hydrophobicity scale.

Statistical sequence analysis & phylogenetic tree. We used the multiple sequence alignment score of the Clustal W algorithm, which is a commonly used progressive multiple sequence alignment method for protein sequences. It can evaluate the similarity between the amino acid sequences. The computer software is available on the website of the Kyoto University Bioinformatics Center, Kyoto, Japan <http://www.genome.jp/tools-bin/clustalw>. The pair wise alignment parameters for the analysis are K-tuple size = 1, window size = 5, gap penalty = 3, and number of top diagonals = 5. The multiple alignment parameters are gap open penalty = 10, gap

extension penalty = 0.05, weight transition = “NO”, hydrophilic residues for proteins = “GPSNDQERK”, hydrophilic gaps = “YES”, and select weight matrix = “BLOSUM (for PROTEIN”.

We compared LDKA peptide (sequence: GLLDLLKLLLKAAG) with LDKA as positive control (completely the same) and lb-AMP2 (sequence QYGRRCCNWGPGRRYCKRWC) as negative control (completely different). Positive control (identical) has the alignment score of 69 and the alignment score of negative control (unlike) is -43. We defined the sequences are similar when alignment score equals or is above 15. AMPs that are similar to LDKA and have alignment score ≥15 are further analyzed and filtered by the source, frog. Then, we built a phylogenetic tree to map these different species of frogs and their geographic location.

We further applied the basic local alignment search tool (BLAST) to analyse the sequences <https://blast.ncbi.nlm.nih.gov/>. This provides the expect value (E-value; E) to check if high similarity scores occur by coincidence or are significant. The E-value is linearly correlated to the effective length of the query (*m*) and the effective length (*n* = total number of bases) of the database; the E-value, defined as $E = m \cdot n \cdot 2^{-S}$, decreases exponentially as the bit score (S) of the match increases. Therefore, lower E-values (and values close to zero) suggest the match is not coincidental.[22]

RESULTS

Simulation-guided peptide design strategy. The goal of this work is to rationalise the *de novo* design of pore-forming peptides by utilising atomic detail structural and dynamic information obtained from MD simulations of spontaneous peptide assembly in membranes. The design strategy is as follows: we start with a simple sequence known to form membrane-inserted helices. Folding-partitioning-assembly simulations are then carried out to obtain the ensemble of membrane-inserted peptide conformations and assemblies this sequence forms in the bilayer. Oligomer-inducing mutations are introduced, guided by analysis of these conformations. Simulation of these mutants provides a new ensemble of structures in the membrane, which in turn is used to inform the next round of mutations. This is continued until a potent pore-forming sequence is obtained. Thus the peptide is designed by stepwise introduction of pore-inducing mutations that are informed by the simulations.

Starting sequence. The design goal is a pore-forming peptide. To avoid introducing any bias with respect to the starting sequence we chose a polyleucine peptide as the initial design template. Leucine is the most abundant amino acid in membrane-active peptides and proteins[4, 16, 23, 24] and polyleucine peptides form stable membrane-spanning helices in lipid bilayers at lengths above 9 residues.[25] We have previously shown that polyleucine peptides partition and fold efficiently into lipid bilayers using unbiased microsecond MD simulations.[13, 21]

As these peptides generally precipitate due to their extreme hydrophobicity we first introduce a single charged lysine as well as a single glycine residue on each peptide terminus to aid solubility. Lysine and glycine were chosen, as together with leucine they are the three most abundant amino acids in short helical AMPs.[4] The basic starting sequence is therefore NH_3^+–GL_nKL_mG–CO_2^-, where *n* & *m* are integers, and the charged amino (NH_3^+–) and carboxylate (–CO_2^-) terminal groups serve to further aid solubility and promote electrostatic peptide-peptide interactions.

The next goal was to find the optimal length of this starting sequence (i.e. the values of *n* and *m*). Natural pore-forming helical AMPs, such as maculatin, are typically between 10-30 residues in length (Figure S2) and retain significant peptide populations bound on the membrane surface, while the remaining peptides insert deeply into the bilayer to form pores.[12, 26, 27] We have previously shown that very similar peptides of the form GL_nRL_nG (*n* = 5-8) bind efficiently to lipid bilayers and continuously flip between surface bound (S) and transmembrane (TM) inserted helical conformations.[16] This allows precise quantification of the relative TM inserted populations from the simulations, with remarkable quantitative accuracy compared to two different experimental measurements.[16] Aiming for a predominantly surface bound peptide, consistent with typical AMPs, we chose the hydrophobic length of our starting sequence to be GL_5KL_6G (i.e. *n* = 5 & *m* = 6), as this lies between the 7±2 % TM-inserted population observed for GL_5RL_5G and the 35±8 % TM population observed for the partitioning-folding simulations of GL_6RL_6G.[9]

Simulation-guided tuning of pore-formation properties. Figure 1A shows that GL_5KL_6G spontaneously inserts into a DMPC:DMPG (3:1) lipid bilayer, which we use here as a simplified model of a bacterial membrane, to form TM helices. In these simulations, a single fully extended peptide is initially placed in bulk solution and allowed to fold and partition freely into and out of the lipid bilayer. As expected from the previous simulations of GL_nRL_nG peptides, TM insertion occurs from membrane surface-bound (S) states and is associated with a free energy change of $\Delta G_{S\rightarrow TM}$ = 0.1±0.4 kcal/mol. For equilibrium simulations, where peptides alternate continuously between S and TM configurations, this can be determined directly using $\Delta G_{S\rightarrow TM} = -RT \ln(p_{TM} / p_S)$, where *R* is the gas constant and *T* is the temperature, and p_{TM} and p_S are the normalised populations of the TM and S states, respectively.

After peptide absorption and folding onto the membrane interface, the system is replicated 16x in the membrane plane, forming a 4x4 matrix, which serves as the starting configuration for a spontaneous pore assembly simulation (Figure 1B). All peptides are initially in the surface state in one interface, but insert over the course of several µs, and start to form various aggregates in the membrane interior. An oligomeric clustering algorithm[12] is used to quantify the numerous TM aggregates formed during the simulations. An oligomer is considered to consist of N peptides if they are all in mutual contact distance, and loosely attached surface (S-state) peptides are omitted to focus on the longer lived TM cores (Figure 1C). Membrane inserted configurations are dominated by V-shaped anti-parallel TM dimers (Figure 1C & Table S2), which are stabilized by hydrogen bonding between the charged terminal amino ($-NH_3^+$) and carboxyl ($-CO_2^-$) groups (Figure 1B). However, no larger pore formation is observed within the 5 µs timeframe of the simulation (Figure S3).

Using this dimer structure to guide mutations that induce channel formation, we first introduce two mutations: L4D and L11K, that convert two leucine residues, located on the same helical face as the central Lys7, into charged amino acids, while keeping the overall charge of the peptide constant at +1. The rationale behind these mutations is to encourage the formation of higher-order oligomeric structures in the membrane by introducing more potential electrostatic interactions between neighbouring peptides.

Pore assembly simulations of this new peptide, NH_3^+-GLLDLLKLLLKLLG-CO_2^- (LDKL), indeed reveal the formation of a water-conducting tetramer within 2 µs (Figures 1C, S1, & Table S2). The simulations reveal that this tetramer is stabilized by two additional inter-helical hydrogen bonds,

compared to GL_5KL_6G. As expected, these are formed by the carboxylate anion of Asp4 interacting with the protonated amino groups of Lys7 and Lys11. This structure is surrounded by a number of surface bound peptides that are in contact with peptides forming a pore (Figure 1A). LDKL has an insertion free energy barrier of $\Delta G_{S\rightarrow TM}$ = 1.2±1.7 kcal/mol. This is higher than GL_5KL_6G ($\Delta G_{S\rightarrow TM}$ = 0.1±0.4 kcal/mol), reflecting the shorter hydrophobic length and two additional charged residues (Table S2). The simulations therefore predict that LDKL is largely surface bound, consistent with the dominant state for most AMPs.

Comparison of the sequence with known AMPs. In the final design step we compare our peptide with known AMP sequences to guide mutations that improve peptide solubility, channel formation efficiency, and pore size. Analysis of the *Antimicrobial Peptide Database* (*APD*) shows that on average ~60 % of amino acids in AMP sequences are hydrophobic (A,F,I,L,V),[23] which is likely to directly influence peptide solubility. We found that many AMPs have statistically lower hydrophobicities in their C-terminal domain. This was calculated by averaging over 6 consecutive terminal residues of all AMPs in the *APD* that have 12 or more amino acids in the sequence using the Wimley & White hydrophobicity scale.[28, 29] To incorporate this into our peptide, two leucines at the C-terminus of LDKL are replaced by two alanines (L12A and L13A), resulting in a new sequence: NH_3^+-GLLDLLKLLLKAAG-CO_2^- (LDKA).

Computational characterization of the designed peptide. Folding-partitioning simulations of a single peptide (Figure 2) predict that LDKA strongly binds and folds onto both DMPC:CHOL:PSM (1:1:1), which serves as a simplified model for a human red blood cell (hRBC) plasma membrane (H), and DMPC:DMPG (3:1), which serves as our very simplified bacterial model membrane (B). These models do not reflect the complexity of a real biological membrane. Instead, they serve to capture the generic physical chemistry of both highly anionic, PG rich, fluid phase bacterial membranes and PC/SPM rich, fluid phase eukaryotic membrane surfaces, respectively. The peptide, which is initially placed in a fully extended conformation in bulk aqueous solvent >15 Å from the membrane spontaneously absorbs and folds onto the interface of the bilayers. No dissociation events are observed over the 5 μs simulation timescale, suggesting strong binding. Once bound, the helicity of monomeric peptides remain ~63 % for both H and B model membranes (Table S3), and the C-terminal –AAG motif remains mostly random coil (Figure S4).

Our previous work shows that the free energy of TM insertion of GL_nRL_nG into zwitterionic POPC bilayers is linearly related to the hydrophobic length of the peptide: $\Delta G_{S\rightarrow TM} = m \cdot n_{Leu} + b$, where n_{Leu} is the number of leucines in the sequence, while m is the increase in insertion free energy per leucine and b is a constant. LDKA, which consists of a total of 14 amino acids including 7 leucines, has a hydrophobic length that is lower than both GL_5RL_5G and GL_6RL_6G. The latter peptides have $\Delta G_{S\rightarrow TM}$ values of 2.1±0.2 kcal/mol (p_{TM} = 7 %) and 0.5±0.3 kcal/mol (p_{TM} = 35 %), respectively.[9] Nevertheless, the computed $\Delta G_{S\rightarrow TM}$ of LDKA is 0.6±0.6 kcal/mol, which is similar to the much more hydrophobic GL_6RL_6G peptide. Figures 1 and 2 show that unlike the GL_nRL_nG family of peptides, which spontaneously TM-insert into the membrane as monomers, LDKA does not insert as a monomer. Instead, LDKA relies on intra-membrane peptide assembly to bury the additional charged residues in a water filled pore.[30]

Structure of the LDKA pores. Spontaneous assembly and pore-formation MD simulation of LDKA shows that this sequence has an increased TM insertion propensity of 34±8 % compared to LDKL (22±8 %), despite having both lower hydrophobicity and helicity. LDKA is 70±3 % helical compared to 90±2 % for LDKL, which is reflected in an increased structural plasticity and experimental solubility (Table S2). The increase in TM-inserted peptide fraction is due to the formation of larger water-conducting oligomeric assemblies. Figure 1C shows that a water-conducting hexamer (Figure S3), which forms within 2 μs remains the dominant intra-membrane oligomeric assembly for this peptide (Figure 1D & Table S2). Like LDKL, the channel-like structures formed by LDKA are surrounded by several surface bound peptides that increase the overall continuous hydrophobic surface facing the membrane, resulting in a better hydrophobic match (Figure 1 & Figure S3). In the oligomeric counting (Figure 1 & 3), we omit more loosely attached surface peptides and consider only sufficiently tilted peptides as part of the oligomeric core (see *Methods*).

LDKA pores are mostly formed by a ring of peptides located in one bilayer leaflet that is matched by a ring of peptides in the other leaflet, resulting in a membrane-spanning pore in which typically 6 peptides form a TM core, and other peptides are in contact at the surface (Figure 1 and 3A). This assembly is stabilized by hydrogen bonds and electrostatic interactions between carboxylate anion of the C-termini and Asp4 with the amino group of the N-termini, as well as Lys7 and Lys11 (Table S2 and Table S3) and hydrophobic contacts (Figure S5).

To investigate the structural stability and plasticity of the LDKA pore-assembly we extended the simulation in the DMPC/DMPG membrane to 15 μs (Figure 3). For this we took the last frame of the 70 ˚C LDKA assembly simulation at 7 μs and continued this simulation for another 8 μs at 50 ˚C. This revealed three equilibrium states: α-helical peptides bound monomerically to the membrane surface, surface assemblies, and transmembrane assemblies. Surface bound peptides in both the upper and bottom bilayer leaflets further extend the hydrophobic length of the pore and help to stabilize the structure (Figure 3A). Figure 3 shows that LDKA peptides flip between S and TM states and are able to cross the hydrophobic core of the bilayer, either via mutual burial of their hydrophilic groups or via shuttling them through the aqueous pore of the oligomeric assembly. Over the 15 μs simulation timescale the dominant structure is made up of a dynamic and disordered stacked pore. This core structure is surrounded by loosely attached peptides bound to the inner and outer membrane interfaces (Figure 3D). The persistence of these large oligomeric pores results in a high level of water leakage across the membrane (Figure 1D), in marked contrast to LDKL or GL_6RL_6G. Dissolution of pores is not observed during the simulations, indicating that TM aggregates are not transient but long-lived low free-energy structures.

Experimental implementation and validation. To validate the simulation-guided peptide design methodology LDKA was synthesized, purified, and functionally characterized *in vitro* using biophysical techniques and biological assays. For the experimental model membranes we chose POPC and POPG instead of DMPC and DMPG to remain above the phase transition temperature when making LUVs via extrusion. The CD and leakage experiments were carried out at room temperature. For the circular dichroism (CD) measurements PSM (extracted from chicken egg; Avanti Polar Lipids, Inc.) was avoided due to the potential presence of protein and impurities that may affect the spectra. All experiments were carried out at room temperature (20 ˚C).

The biological activity of pore-forming AMPs depends crucially on their ability to bind strongly to lipid bilayers without precipitating out of aqueous solution.[11, 31] Figure 4A shows CD spectra of LDKA in aqueous solution as well as in the presence of POPC:POPG (1:1) and POPC:CHOL (7:3) large unilamellar vesicles (LUVs) at pH = 7. The results confirm that LDKA is indeed soluble (< 0.1 mg/mL in 1x pH 7.4 DPBS) and partially helical in solution, and that it spontaneously binds to both model membranes, increasing peptide helicity from 45 % to 78 %. The CD-measured helical fraction of membrane-bound LDKA peptides agrees with the equilibrium LDKA assembly simulations at 50 ˚C, which is 78±2 % (Table S2).

LDKA pore-formation was quantified using a membrane leakage assay. This assay measures the amount of liposome-encapsulated dye released into the external solution by pores formed in the liposomal membrane or by membrane disruption after addition of LDKA. Figure 4B shows that LDKA results in fluorescent dye leakage from both POPC:POPG (1:1) and POPC:CHOL:PSM (1:1:1), serving as simplified bacterial and hRBC model membranes, respectively. Leakage is higher for POPC:POPG vesicles with >50 % release of ANTS/DPX (350 Da) dyes even at very low peptide-to-lipid (P/L) ratios of 1/1000. Variation of the molecular weight of the encapsulated dye (350 Da ANTS/DPX fluorophores and 10-kDa dextrans) allows estimating the size of the membrane pore. While leakage from POPC:POPG vesicles is ~100 % at P/L > 1/50, irrespective of the dye size, leakage of the larger 10-kDa dextrans is only a fraction of the small dye leakage for P/L < 1/100. This suggests that LDKA forms an ensemble of pores of different sizes, with smaller pores dominating at low P/L ratios, which is consistent with the structures observed in the channel assembly simulations (Figure 1 & Figure 3). A similar ensemble of structurally different membrane pores has recently been demonstrated for the naturally occurring AMP maculatin.[12]

Antimicrobial and hemolytic activity. Biological activity of LDKA is determined using bacterial minimum inhibitory concentration (MIC) and hemolysis assays. The results are summarized in Figure 4C, demonstrating that LDKA has antibacterial activity against both gram-positive (*S. aureus*) and gram-negative (*E. coli* and *P. aeruginosa*) bacteria in nutrient-rich medium, with MICs in the 10-70 µM range, which are typical for naturally occurring AMPs.[32-37] We note that both *P. aeruginosa* and *S. aureus* are resistant to some AMPs. Figure 4D shows that similar to indolicidin, a short 13-residue natural bovine AMP, LDKA causes minimal damage (< 5% fractional hemolysis) to hRBCs below concentrations of 20 µM. Above this concentration, hemolysis is higher than that induced by indolicidin. These results suggest that LDKA has sufficient antimicrobial potency and selectivity that it could be a good starting point for optimization towards the development of potential pharmaceutical applications.

LDKA comparison with naturally occurring AMPs. In order to investigate how similar the designed LDKA sequence is to naturally evolved AMPs we performed a comparison of the functional motif with 2,619 AMP sequences currently deposited in the APD.[4] Figure S6A and Table S5 show that 224 AMPs share the LDKA motif: shhXhhXhhhXhhs, where s is a small residue (G,A,S), X is a charged residue (D,E,K,R), and h is a hydrophobic residue (A,F,I,L,V). Many sequences have N-terminal GhhD or ShhK motifs, and contain KhhhK motifs in the middle of their sequences (Table S4),[4, 23] and several short AMPs have very similar motifs, e.g. Hylaseptin P1,[35] Hylain 2,[37] DFTamP1,[23] Frenatin 2,[32] Aurein 1.1,[38] Dahlein 1.1,[34] Fallaxidin 3.1,[36] and Uperin 7.1.[33]

Further analysis revealed that 199 of the 224 AMPs with the LDKA motif are from frogs, corresponding to ~22 % of all frog AMPs currently deposited in the APD (Figure S6B). Interestingly, these frogs are distributed all over the world (Figure S6C), suggesting that this is an optimized pore-forming sequence motif that evolved in a common ancestor of current frogs before the break-up of the Pangea supercontinent in the middle of the Jurassic around 175 million years ago.[39]

The chance of the shhXhhXhhhXhhs motif to develop by chance is $7 \cdot 10^{-10}$, by comparison, the chance of any unique 14 residue sequence is $6 \cdot 10^{-19}$. We performed a BLAST analysis to obtain a measure of how likely LDKA (GLLDLLKLLLKAAG) is to occur by coincidence and found 103 top BLAST hits (> 0.7 identical value; Figure S7A). The analysis showed that more than half of them have an E-value of 116, which suggests low significance. 9 sequences suggested a significant match with E-values ≤ 10. While this suggests these protein sequences are highly related, these sequences are all from bacteria and generally form part of larger proteins (Figure S7B and Table S6) and are therefore unlikely to have a common evolutionary ancestor.[22] Further BLAST analysis of 8 LDKA-like frog AMPs (Figure S7C) shows that three out of these eight peptides have more sequences with E-value scores ≤ 10 than LDKA ($N_{E\text{-}value \leq 10} = 9$), namely: Hylain 2 ($N_{E\text{-}value \leq 10} = 100$), Frenatin 2 ($N_{E\text{-}value \leq 10} = 59$), and Aurein 1.1 ($N_{E\text{-}value \leq 10} = 19$). Thus the fact that the shhXhhXhhhXhhs motif occurs as an AMP in many evolutionarily distant frogs does not necessarily suggest that it is the result of a common ancestor, as it also occurs in evolutionarily unrelated species and as part of larger proteins. This is probably due to this motif encompassing over 10^{10} possible members, which makes it quite likely to have formed evolutionarily independently.

DISCUSSION

The design of novel AMPs has been hindered by our patchy understanding of the molecular mechanisms underlying their antimicrobial activity. This paucity of mechanistic information is chiefly due to the absence of high-resolution structural data, as resolving the complex and dynamic interactions of flexible and structurally dynamic AMPs with their cellular targets is extremely challenging. Consequently, previous design approaches have largely been based on bioinformatic analyses of known AMP sequences, or empirical designs often based on helical wheel projection in the case of pore-forming AMPs. Specifically, these approaches has been used to design AMP sequences using database-guided *in silico* design,[23, 24] experimental combinatorial library screening,[7-11, 40] and rational engineering.[41-43] Functional refinement of these designs can be achieved via high-throughput library-screening approaches.[44] While these approaches have been successfully employed to discover potent new AMPs, the lack of detailed structural information has so far precluded a more rational design process.

Long timescale equilibrium MD simulations now offer an alternative avenue towards elucidating the interactions of pore-forming AMPs with fluid lipid bilayers. Here we show that the information obtained from these simulations can be utilized to design pore-forming AMPs. In particular, the simulations provide details of the structural interactions of specific sequences in the context of a fluid lipid bilayer. Analysis of these interactions is used to propose mutations that favor pore formation and stabilisation.

In the present case, the simulations revealed that LDKA forms an unusual pore formed by two stacked rings of peptides, one in each bilayer leaflet, surrounding a central water-filled

channel. This largely disordered structure is stabilized by intermolecular salt bridges and hydrogen-bonds that continuously form and disband in the membrane reflecting a mechanism previously proposed for δ-lysin based on experimental evidence.[45]

Results obtained from pore-assembly simulations can be experimentally validated. In general MD simulation can be validated via a number of experimental measurements. For example, the average populations of surface-bound and membrane inserted conformations can be obtained via oriented CD,[15] while induced leakage can be checked via vesicle leakage assays, where the size of the molecule leaking out of peptide-induced pores can provide information on the pore size.[12] Binding as well as folding-partitioning can be determined by measuring the changes to tryptophan-fluorescence or CD spectra upon titration of LUVs. While these methods do not provide the wealth of information contained in the MD simulations, they can serve as validating and calibrating tools. In the present study we apply both CD and leakage assays to computationally characterise the final LDKA design.

An interesting characteristic of LDKA is that a single helical peptide is not long enough to span the entire hydrophobic core thickness of a bacterial membrane. Nevertheless LDKA is as efficient as the 26-residue designed membrane-active peptide Melp5 (a synthetically evolved gain-of-function variant of melittin) causing similar leakage of ANTS/DPX dyes from both POPC and POPG vesicles[7, 8]. Comparison of the *in vitro* activity further revealed that LDKA is potent against *E. coli* (MIC = 35±9 μM) and *S. aureus* (MIC = 10±0 μM) and comparable to Melp5 ($MIC_{E.coli}$ = 4 μM and $MIC_{S.aureus}$ > 40 μM) and melittin ($MIC_{E.coli}$ = 3 μM and $MIC_{S.aureus}$ = 4 μM). LDKA also has lower hemolytic activity (HC_{50} = 63±8 μM), compared to both Melp5 and melittin (HC_{50} < 5 μM)[8]. Unfortunately, comparing potencies of different AMPs provides precious little information regarding their mechanisms of activity, nor does it provide information on the pore size or structures. Too little is currently known to correlate pore size of the membrane structure with *in-vitro* activity, and pore-forming AMPs may have other non-membrane targets that may act as confounders in such an analysis. For example, empirically designed pepides, like the pHD peptides[10, 11] and macrolittins[9] are extraordinarily potent at permeabilising membranes; more potent than all known AMPs and most other membrane permeabilizing peptides, but their unique and fascinating mechanism of action remains unknown and is unlikely to be similar to LDKA. Antimicrobial peptides, even those with excellent biological activity, never have such extraordinary potency against lipid vesicles.

In general, the LDKA peptide is most similar in its mechanistic properties to pore-forming AMPs such as maculatin. These types of membrane-active peptides form ensembles of well defined pores, which are amenable to structural characterisation using atomic detail MD simulations as presented here. In contrast, peptides that cause large-scale disruption of membranes require simulations on a much larger scale, which will require much larger systems, beyond the reach of current hardware.

CONCLUSION

The results presented here show that atomic detail folding-assembly simulations are now suitably evolved to guide the *de novo* design of pore-forming membrane-active peptides. Starting from a hydrophobic polyleucine template sequence, the design process arrives at a potent functional motif, that is related to a similar motif occurring naturally in frogs. In addition, the

simulations reveal how this sequence motif binds, folds, oligomerizes, and forms pores in model membranes.

Such simulations have not been previously attempted due to the belief that the timescales to observe spontaneous channel formation is not reachable yet. However, the present study demonstrates that the effect of small mutations on the prediction of channel equilibria and structures can now be directly quantified via all-atomistic MD simulations, as shown previously for maculatin.[12] This establishes equilibrium MD simulations as a tool to investigate sequence dependencies that can be very difficult to understand, due to the extreme heterogeneity of the various channel aggregates formed. The small alterations of the 3 sequences devised in this work lead to very different channel equilibria, with a huge difference in membrane poration and leakage, all of which would be very challenging to predict via bioinformatic sequence-analysis tools. Both the pore-forming properties as well as antimicrobial activity of the designed peptide can then be confirmed experimentally, validating the design methodology presented. The heterogeneous and transient nature of the many pore-like aggregates formed by these peptides demonstrate that *in silico* quantification of poration via MD simulations, as shown here, may be the key element in the future development of membrane permeabilizing sequences. This opens the door to rational design and optimization of other membrane-active peptides.

ASSOCIATED CONTENT

Supporting Information

Secondary structure of LDKA peptide against elevated temperature, peptide length analysis of the APD, water conducting structures, residue helicity of each LDKA and LDKL with DMPC/DMPG bilayer, contact analysis of the LDKA channel, comparison of LDKA with naturally occurring AMPs, homology analysis of the LDKA sequence using BLAST, and summary of the simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Correspondence to: Charles H. Chen (charles.chen@kcl.ac.uk), Jakob P. Ulmschneider (jakob@sjtu.edu.cn), and Martin B. Ulmschneider (martin.ulmschneider@kcl.ac.uk).

NOTES

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

We thank Guangshun Wang at the University of Nebraska Medical Center for providing the raw data of the antimicrobial peptide database. Simulations were made possible by the MARCC supercomputer facility at Johns Hopkins University.

REFERENCES

1. Zasloff, M., Magainins, a class of antimicrobial peptides from Xenopus skin: isolation, characterization of two active forms, and partial cDNA sequence of a precursor. *Proc Natl Acad Sci U S A* **1987,** *84* (15), 5449-53.
2. Lehrer, R. I.; Barton, A.; Daher, K. A.; Harwig, S. S.; Ganz, T.; Selsted, M. E., Interaction of human defensins with Escherichia coli. Mechanism of bactericidal activity. *J Clin Invest* **1989,** *84* (2), 553-61.
3. Yeaman, M. R.; Yount, N. Y., Mechanisms of antimicrobial peptide action and resistance. *Pharmacol Rev* **2003,** *55* (1), 27-55.
4. Wang, G.; Li, X.; Wang, Z., APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res* **2016,** *44* (D1), D1087-93.
5. Wimley, W. C.; Hristova, K., Antimicrobial peptides: successes, challenges and unanswered questions. *J Membr Biol* **2011,** *239* (1-2), 27-34.
6. Wiedman, G.; Fuselier, T.; He, J.; Searson, P. C.; Hristova, K.; Wimley, W. C., Highly efficient macromolecule-sized poration of lipid bilayers by a synthetically evolved peptide. *J Am Chem Soc* **2014,** *136* (12), 4724-31.
7. Krauson, A. J.; He, J.; Wimley, W. C., Gain-of-function analogues of the pore-forming peptide melittin selected by orthogonal high-throughput screening. *J Am Chem Soc* **2012,** *134* (30), 12732-41.
8. Krauson, A. J.; Hall, O. M.; Fuselier, T.; Starr, C. G.; Kauffman, W. B.; Wimley, W. C., Conformational Fine-Tuning of Pore-Forming Peptide Potency and Selectivity. *J Am Chem Soc* **2015,** *137* (51), 16144-52.
9. Li, S.; Kim, S. Y.; Pittman, A. E.; King, G. M.; Wimley, W. C.; Hristova, K., Potent Macromolecule-Sized Poration of Lipid Bilayers by the Macrolittins, A Synthetically Evolved Family of Pore-Forming Peptides. *J Am Chem Soc* **2018**, *140* (20), 6441-7.
10. Wiedman, G.; Wimley, W. C.; Hristova, K., Testing the limits of rational design by engineering pH sensitivity into membrane-active peptides. *Biochim Biophys Acta* **2015,** *1848* (4), 951-7.
11. Wiedman, G.; Kim, S. Y.; Zapata-Mercado, E.; Wimley, W. C.; Hristova, K., pH-Triggered, Macromolecule-Sized Poration of Lipid Bilayers by Synthetically Evolved Peptides. *J Am Chem Soc* **2017,** *139* (2), 937-945.
12. Wang, Y.; Chen, C. H.; Hu, D.; Ulmschneider, M. B.; Ulmschneider, J. P., Spontaneous formation of structurally diverse membrane channel architectures from a single antimicrobial peptide. *Nat Commun* **2016,** *7*, 13535.
13. Ulmschneider, J. P.; Smith, J. C.; White, S. H.; Ulmschneider, M. B., In silico partitioning and transmembrane insertion of hydrophobic peptides under equilibrium conditions. *J Am Chem Soc* **2011,** *133* (39), 15487-95.
14. Andersson, M.; Ulmschneider, J. P.; Ulmschneider, M. B.; White, S. H., Conformational states of melittin at a bilayer interface. *Biophys J* **2013,** *104* (6), L12-4.
15. Chen, C. H.; Wiedman, G.; Khan, A.; Ulmschneider, M. B., Absorption and folding of melittin onto lipid bilayer membranes via unbiased atomic detail microsecond molecular dynamics simulation. *Biochim Biophys Acta* **2014,** *1838* (9), 2243-9.

16. Ulmschneider, M. B.; Ulmschneider, J. P.; Schiller, N.; Wallace, B. A.; von Heijne, G.; White, S. H., Spontaneous transmembrane helix insertion thermodynamically mimics translocon-guided insertion. *Nat Commun* **2014,** *5*, 4863.
17. Wang, Y.; Zhao, T.; Wei, D.; Strandberg, E.; Ulrich, A. S.; Ulmschneider, J. P., How reliable are molecular dynamics simulations of membrane active antimicrobial peptides? *Biochim Biophys Acta* **2014,** *1838* (9), 2280-8.
18. Chen, C. H.; Khan, A.; Huang, J. J.; Ulmschneider, M. B., Mechanisms of Membrane Pore Formation by Amyloidogenic Peptides in Amyotrophic Lateral Sclerosis. *Chemistry* **2016,** *22* (29), 9958-61.
19. Ulmschneider, J. P.; Ulmschneider, M. B., Molecular Dynamics Simulations Are Redefining Our View of Peptides Interacting with Biological Membranes. *Acc Chem Res* **2018,** *51* (5), 1106-1116.
20. Ulmschneider, M. B.; Ulmschneider, J. P., Folding Peptides into Lipid Bilayer Membranes. *J Chem Theory Comput* **2008,** *4* (11), 1807-9.
21. Ulmschneider, M. B.; Doux, J. P.; Killian, J. A.; Smith, J. C.; Ulmschneider, J. P., Mechanism and kinetics of peptide partitioning into membranes from all-atom simulations of thermostable peptides. *J Am Chem Soc* **2010,** *132* (10), 3452-60.
22. Ye, J.; McGinnis, S.; Madden, T. L., BLAST: improvements for better sequence analysis. *Nucleic Acids Res* **2006,** *34* (Web Server issue), W6-9.
23. Mishra, B.; Wang, G., Ab initio design of potent anti-MRSA peptides based on database filtering technology. *J Am Chem Soc* **2012,** *134* (30), 12426-9.
24. Porto, W. F.; Irazazabal, L.; Alves, E. S. F.; Ribeiro, S. M.; Matos, C. O.; Pires, Á.; Fensterseifer, I. C. M.; Miranda, V. J.; Haney, E. F.; Humblot, V.; Torres, M. D. T.; Hancock, R. E. W.; Liao, L. M.; Ladram, A.; Lu, T. K.; de la Fuente-Nunez, C.; Franco, O. L., In silico optimization of a guava antimicrobial peptide enables combinatorial exploration for peptide design. *Nat Commun* **2018,** *9* (1), 1490.
25. Jaud, S.; Fernández-Vidal, M.; Nilsson, I.; Meindl-Beinker, N. M.; Hübner, N. C.; Tobias, D. J.; von Heijne, G.; White, S. H., Insertion of short transmembrane helices by the Sec61 translocon. *Proc Natl Acad Sci U S A* **2009,** *106* (28), 11588-93.
26. Parton, D. L.; Akhmatskaya, E. V.; Sansom, M. S., Multiscale simulations of the antimicrobial peptide maculatin 1.1: water permeation through disordered aggregates. *J Phys Chem B* **2012,** *116* (29), 8485-93.
27. Fernandez, D. I.; Lee, T. H.; Sani, M. A.; Aguilar, M. I.; Separovic, F., Proline facilitates membrane insertion of the antimicrobial peptide maculatin 1.1 via surface indentation and subsequent lipid disordering. *Biophys J* **2013,** *104* (7), 1495-1507.
28. Wimley, W. C.; White, S. H., Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* **1996,** *3* (10), 842-8.
29. White, S. H.; Wimley, W. C., Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* **1999,** *28*, 319-65.
30. Walther, T. H.; Ulrich, A. S., Transmembrane helix assembly and the role of salt bridges. *Curr Opin Struct Biol* **2014,** *27*, 63-8.
31. Farrand, A. J.; Hotze, E. M.; Sato, T. K.; Wade, K. R.; Wimley, W. C.; Johnson, A. E.; Tweten, R. K., The Cholesterol-dependent Cytolysin Membrane-binding Interface Discriminates Lipid

Environments of Cholesterol to Support β-Barrel Pore Insertion. *J Biol Chem* **2015,** *290* (29), 17733-44.
32. Raftery, M. J.; Waugh, R. J.; Bowie, J. H.; Wallace, J. C.; Tyler, M. J., The structures of the frenatin peptides from the skin secretion of the giant tree frog Litoria infrafrenata. *J Pept Sci* **1996,** *2* (2), 117-24.
33. Steinborner, S. T.; Bowie, J. H.; Tyler, M. J.; Wallace, J. C., An Unusual Combination of Peptides from the Skin Glands of Ewing's Tree Frog, Litoria ewingi. Sequence Determination and Antimicrobial Activity. *Australian Journal of Chemistry* **1997,** *50* (9), 889-894.
34. Wegener, K. L.; Brinkworth, C. S.; Bowie, J. H.; Wallace, J. C.; Tyler, M. J., Bioactive dahlein peptides from the skin secretions of the Australian aquatic frog Litoria dahlii: sequence determination by electrospray mass spectrometry. *Rapid Commun Mass Spectrom* **2001,** *15* (18), 1726-34.
35. Prates, M. V.; Sforça, M. L.; Regis, W. C.; Leite, J. R.; Silva, L. P.; Pertinhez, T. A.; Araújo, A. L.; Azevedo, R. B.; Spisni, A.; Bloch, C., The NMR-derived solution structure of a new cationic antimicrobial peptide from the skin secretion of the anuran Hyla punctata. *J Biol Chem* **2004,** *279* (13), 13018-26.
36. Jackway, R. J.; Bowie, J. H.; Bilusich, D.; Musgrave, I. F.; Surinya-Johnson, K. H.; Tyler, M. J.; Eichinger, P. C., The fallaxidin peptides from the skin secretion of the Eastern Dwarf Tree Frog Litoria fallax. Sequence determination by positive and negative ion electrospray mass spectrometry: antimicrobial activity and cDNA cloning of the fallaxidins. *Rapid Commun Mass Spectrom* **2008,** *22* (20), 3207-16.
37. Wu, J.; Liu, H.; Yang, H.; Yu, H.; You, D.; Ma, Y.; Ye, H.; Lai, R., Proteomic analysis of skin defensive factors of tree frog Hyla simplex. *J Proteome Res* **2011,** *10* (9), 4230-40.
38. Rozek, T.; Wegener, K. L.; Bowie, J. H.; Olver, I. N.; Carver, J. A.; Wallace, J. C.; Tyler, M. J., The antibiotic and anticancer active aurein peptides from the Australian Bell Frogs Litoria aurea and Litoria raniformis the solution structure of aurein 1.2. *Eur J Biochem* **2000,** *267* (17), 5330-41.
39. San Mauro, D.; Vences, M.; Alcobendas, M.; Zardoya, R.; Meyer, A., Initial diversification of living amphibians predated the breakup of Pangaea. *Am Nat* **2005,** *165* (5), 590-9.
40. Fuselier, T.; Wimley, W. C., Spontaneous Membrane Translocating Peptides: The Role of Leucine-Arginine Consensus Motifs. *Biophys J* **2017,** *113* (4), 835-846.
41. Sani, M. A.; Lee, T. H.; Aguilar, M. I.; Separovic, F., Proline-15 creates an amphipathic wedge in maculatin 1.1 peptides that drives lipid membrane disruption. *Biochim Biophys Acta* **2015,** *1848* (10 Pt A), 2277-89.
42. Torres, M. D. T.; Pedron, C. N.; Higashikuni, Y.; Kramer, R. M.; Cardoso, M. H.; Oshiro, K. G. N.; Franco, O. L.; Silva Junior, P. I.; Silva, F. D.; Oliveira Junior, V. X.; Lu, T. K.; de la Fuente-Nunez, C., Structure-function-guided exploration of the antimicrobial peptide polybia-CP identifies activity determinants and generates synthetic therapeutic candidates. *Commun Biol* **2018,** *1*, 221.
43. Kauffman, W. B.; Guha, S.; Wimley, W. C., Synthetic molecular evolution of hybrid cell penetrating peptides. *Nat Commun* **2018,** *9* (1), 2568.
44. Lau, J. L.; Dunn, M. K., Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg Med Chem* **2018,** *26* (10), 2700-2707.
45. Pokorny, A.; Birkbeck, T. H.; Almeida, P. F., Mechanism and kinetics of delta-lysin interaction with phospholipid vesicles. *Biochemistry* **2002,** *41* (36), 11044-56.

Figure 1

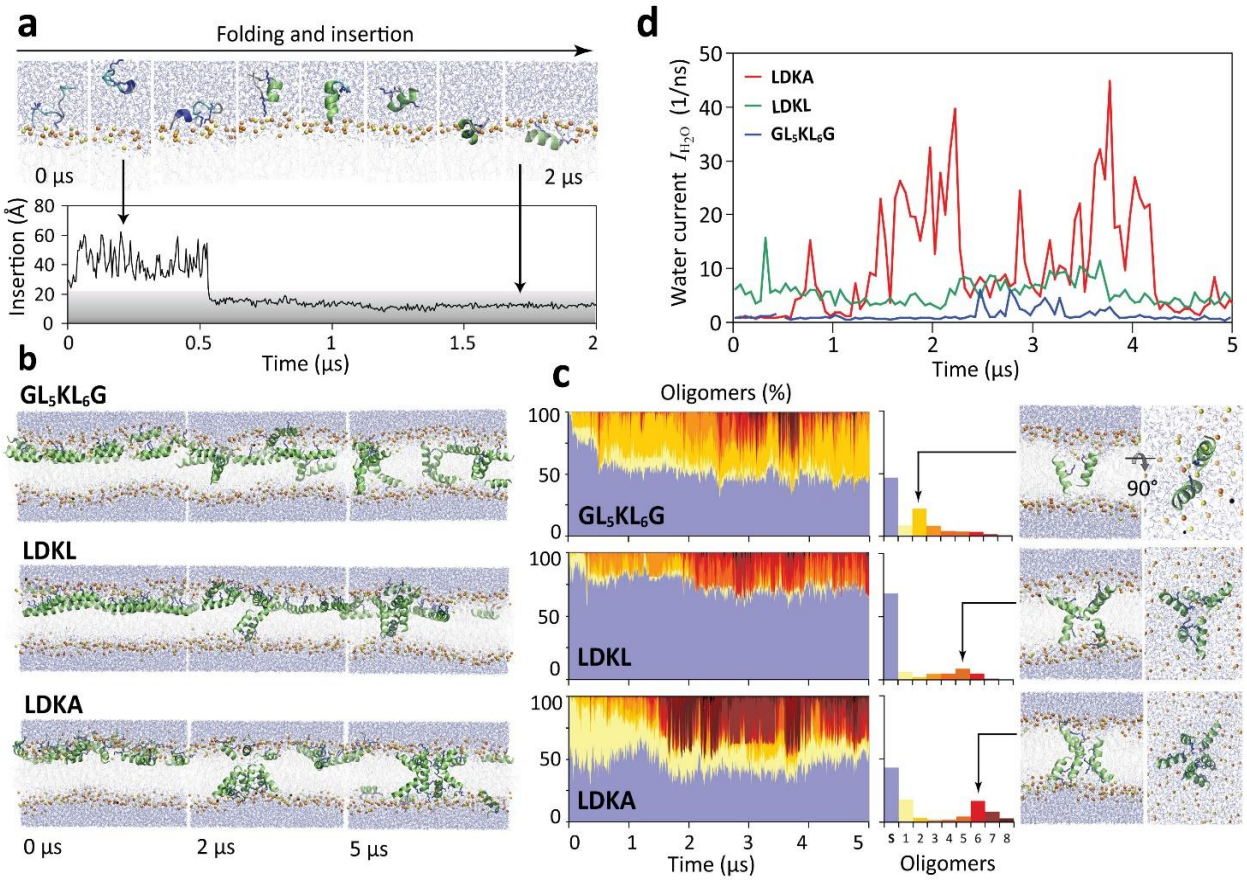

Figure 1 | Unbiased multi-microsecond MD simulations reveal the spontaneous assembly of designed sequences into oligomeric structural ensembles in the membrane. a, Starting from an initially unfolded structure, single peptides spontaneously fold and insert into the membrane into a surface bound S-state, (lower panel: peptide center of mass insertion depth in the membrane, 0 Å = membrane center). **b,** The bound S-state system is multiplexed to model the adsorption of numerous S-state peptides into the outer leaflet. Cooperative TM insertion and oligomerization is observed for GL_5KL_6G, LDKL, and LDKA. **c,** Oligomer analysis reveals the predominant multimers in these systems are dimers (for GL_5KL_6G), pentamers (for LDKL), and hexamers (for LDKA), respectively (color coded, bar chart on the right). Octamers were the largest channels observed. Aggregates of more than 8 peptides exist, but consist of several lower order assemblies in lateral contact, and occur infrequently. **d,** The overall bilayer patch water conductance reveals that larger pores lead to a larger leakage over the course of the simulations.

Figure 2

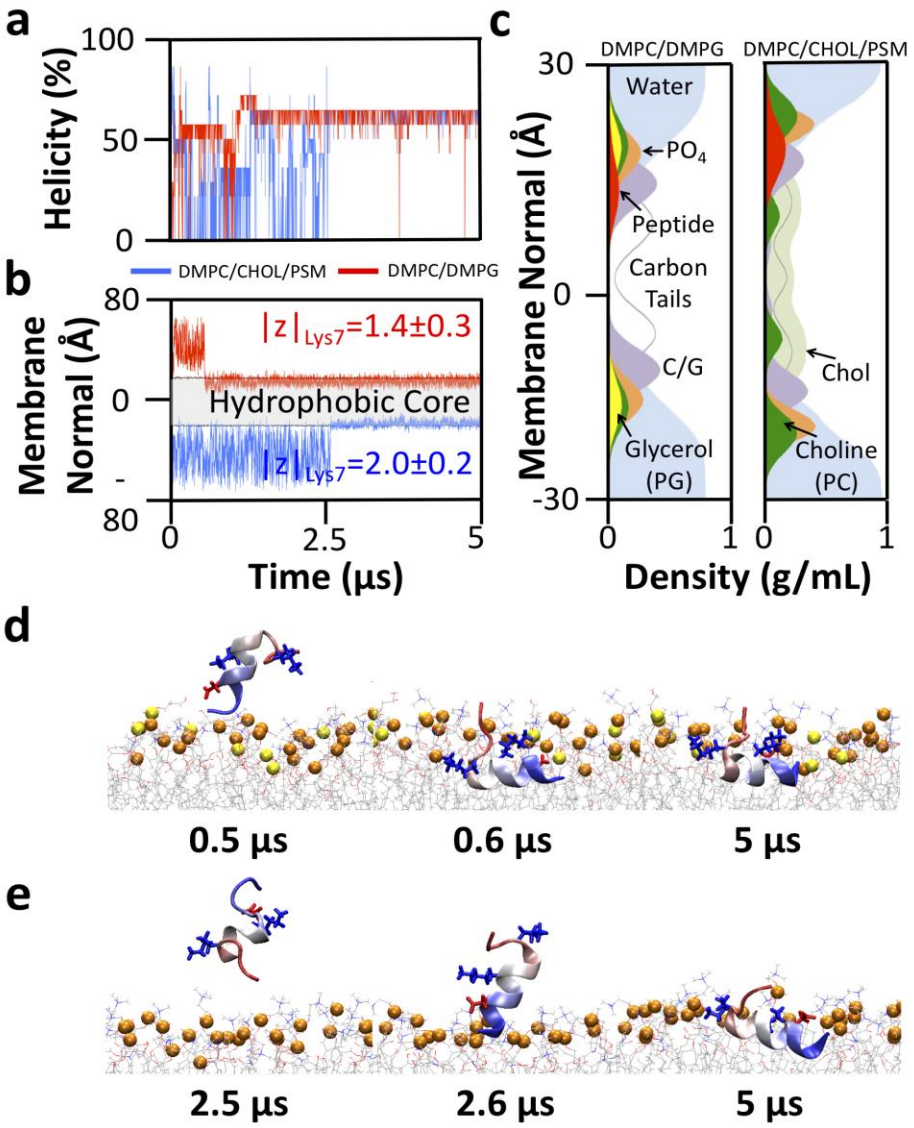

Figure 2. | Spontaneous folding of LDKA with DMPC/CHOL/PSM bilayer and DMPC/DMPG bilayer at 70 °C. a, Corresponding evolution of the helicity, and **b,** the insertion depth of Lys7 of LDKA in membrane lipids. **c,** The density cross-section profile of the bilayer shows that the folded S state LDKA peptide is buried below the water interface, and has overlap with the phosphate headgroups (PC & PG) and carbonyl-glycerol (C/G) group. **d,** Snapshots of LDKA binding and folding with DMPC/DMPG bilayer (ratio of 3:1; simplified bacterial model membrane) and **e,** DMPC/CHOL/PSM (ratio of = 1:1:1; simplified human red blood cell model membrane). Ribbons colored blue to red from N- to C-terminus, and positive and charged sidechains are shown as blue and red, respectively. Orange sphere represents the phosphate headgroup of the zwitterionic lipids, and yellow sphere shows the phosphate headgroup of the anionic lipids.

Figure 3

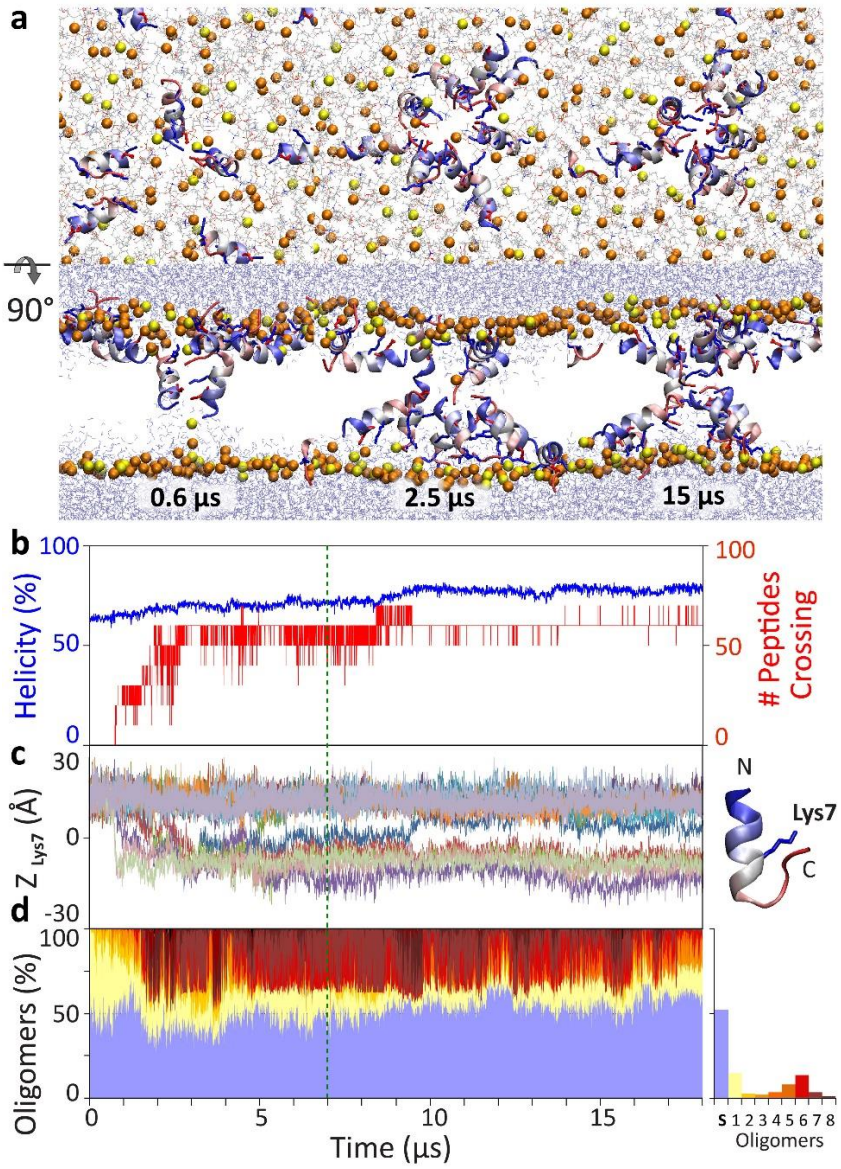

Figure 3 | LDKA hexamer assembly and stability. a, Snapshots of top and side views of peptides assembling on the surface of the membrane, inserting as dimer into the membrane, and forming a pore. The peptide backbone is color coded from blue (N-terminus) to red (C-terminus). Positive and negative-charged residues are shown in blue and red, respectively. Orange spheres and yellow spheres represent the phosphate headgroup of DMPC and DMPG, respectively. **b,** Helical fraction and number of LDKA peptides crossing the membrane. **c,** Distribution of Lys7 z-position as a label to track all LDKA peptides. The double stacking of the pore is clearly visible, with some peptides in the upper and others in the lower leaflet. **d,** TM multimer analysis shows that the hexamer is the dominant conformation over the 18 µs timescale of the simulation. The dashed green line shows where the simulation temperature was reduced from 70 to 50 ˚C.

Figure 4

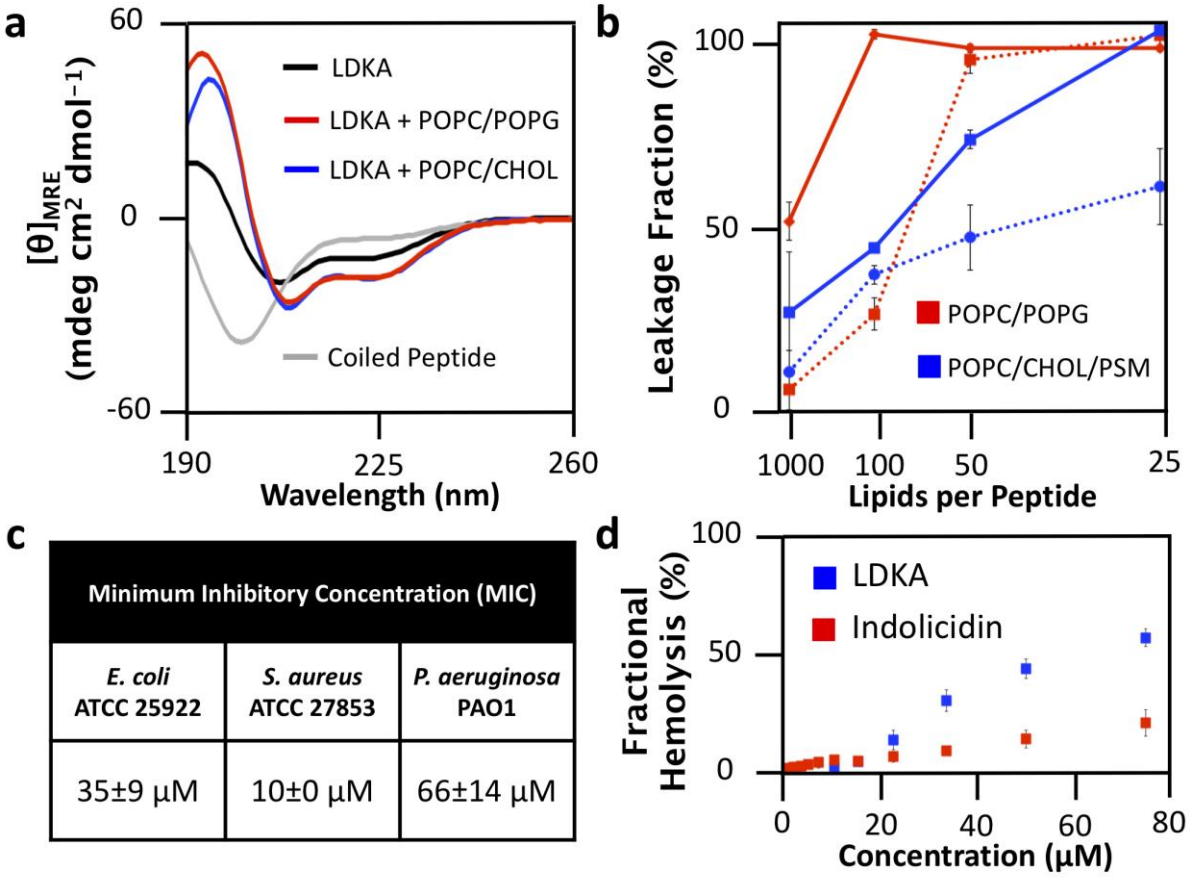

Figure 4 | *In vitro* experiments show LDKA is a helical membrane-active antimicrobial peptide. **a,** Circular dichroism spectra of 50 µM LDKA in buffer (black), LDKA in the presence of POPC/CHOL vesicle (red, P:L = 1:12.5; POPC:CHOL = 7:3; simplified human red blood cell model membrane), and LDKA in the presence of POPC/POPG vesicle (blue, P:L = 1:12.5; POPC:POPG = 3:1; simplified bacteria model membrane) in 10 mM phosphate buffer at pH 7. The peptide is partially helical in solution and folds onto both bilayer models. A coiled peptide spectra is shown for reference (grey; reference: Hylaseptin P1). **b,** Dye leakage assay of LDKA with POPC/CHOL/PSM vesicle (blue; POPC:CHOL:PSM = 1:1:1; simplified human red blood cell model membrane) and POPC/POPG vesicle (red; POPC:POPG = 3:1; simplified bacteria model membrane). Solid lines mean that ANTS/DPX (M.W. = 400) is the trapped fluorescent marker in the vesicle, while dashed lines show data for trapped 10-kDa Dextran. **c,** Bacterial minimum inhibition concentration (MIC) shows LDKA is a powerful AMP. Initial bacterial cell density was prepared with ~3 x 10^5 CFU/mL. LDKA was co-incubated with the bacteria in Lysogeny broth at 37 °C for 12 hr. The wavelength of optical density was determined at 600 nm. **d,** Hemolysis assay of titrated LDKA with red blood cell. Indolicidin, an AMP with 13 amino acids with coiled structure, was used as a reference.

TOC Graphic